



Using `stargazer` to report regression output and descriptive statistics in R (for *non-LaTeX* users)

(v1.0 draft)

Oscar Torres-Reyna

otorres@princeton.edu



May 2014

 <http://dss.princeton.edu/training/>

Introduction

As anything with R, there are many ways of exporting output into nice tables (but mostly for *LaTeX* users). Some packages are: `apsrtable`, `xtable`, `texreg`, `memisc`, `outreg` ...and counting.

At the moment, the “new kid on the block” is `stargazer`. Released by Marek Hlavac on March 3rd, 2014, version 5.0 offers a very nice, smart, and easy-to-use alternative to non-*LaTeX* users, in particular, the ability to import editable tables into a Word document.

This presentation will show some of the options `stargazer` offers, the contents are based on the documentation from the package available in the following links:

- <http://cran.r-project.org/web/packages/stargazer/stargazer.pdf>
- <http://cran.r-project.org/web/packages/stargazer/vignettes/stargazer.pdf>

The default setting produces *LaTeX* code, the additional alternatives are:

- Output as text, which allows a quick view of results
- Output as html, which produce editable tables for Word documents.

For a quick view of the results use the “text” format option

OUTPUT IN TEXT FORMAT

Descriptive statistics: in text format

```
mydata <- mtcars
install.packages("stargazer") #Use this to install it, do this only once
library(stargazer)
stargazer(mydata, type = "text", title="Descriptive statistics", digits=1, out="table1.txt")
```

stargazer will automatically recognize the type of object, and will produce the appropriate output. In the case of data frames, it will display summary statistics .

```
Descriptive statistics
=====
Statistic N Mean St. Dev. Min Max
mpg      32 20.1   6.0   10.4 33.9
cyl      32  6.2   1.8    4    8
disp     32 230.7 123.9  71.1 472.0
hp       32 146.7  68.6   52   335
drat     32  3.6   0.5    2.8  4.9
wt       32  3.2   1.0    1.5  5.4
qsec     32 17.8   1.8   14.5 22.9
vs       32  0.4   0.5    0    1
am       32  0.4   0.5    0    1
gear     32  3.7   0.7    3    5
carb     32  2.8   1.6    1    8
-----
```

↑
The table will be saved in the working directory with whatever name you write in the out option. You can open this file with any word processor

```
# Same output, transposed (variables in columns)
stargazer(mydata, type = "text", title="Descriptive statistics", digits=1, out="table1.txt", flip=TRUE)
```

```
Descriptive statistics
=====
Statistic mpg cyl disp hp drat wt qsec vs am gear carb
N          32  32  32   32  32  32  32  32  32  32  32
Mean       20.1 6.2 230.7 146.7 3.6 3.2 17.8 0.4 0.4 3.7 2.8
St. Dev.   6.0 1.8 123.9 68.6 0.5 1.0 1.8 0.5 0.5 0.7 1.6
Min        10.4 4 71.1 52 2.8 1.5 14.5 0 0 3 1
Max        33.9 8 472.0 335 4.9 5.4 22.9 1 1 5 8
-----
```

↑
Use this option if you want the variables in columns

Descriptive statistics: in text format, replacing variable names with labels

The table will be saved in the working directory with whatever name you write in the `out` option. You can open this file with any word processor

```
mydata <- mtcars
install.packages("stargazer") #Use this to install it, do this only once
library(stargazer)
stargazer(mydata, type = "text", title="Descriptive statistics", digits=1, out="table1.txt",
          covariate.labels=c("Miles/(US)gallon","No. of cylinders","Displacement (cu.in.)",
                             "Gross horsepower","Rear axle ratio","Weight (lb/1000)",
                             "1/4 mile time","V/S","Transmission (0=auto, 1=manual)",
                             "Number of forward gears","Number of carburetors"))
```

Use the option `covariate.labels` to replace variable names with variable labels. Must be in same order as in the dataset.

Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Miles/(US)gallon	32	20.1	6.0	10.4	33.9
No. of cylinders	32	6.2	1.8	4	8
Displacement (cu.in.)	32	230.7	123.9	71.1	472.0
Gross horsepower	32	146.7	68.6	52	335
Rear axle ratio	32	3.6	0.5	2.8	4.9
Weight (lb/1000)	32	3.2	1.0	1.5	5.4
1/4 mile time	32	17.8	1.8	14.5	22.9
V/S	32	0.4	0.5	0	1
Transmission (0=auto, 1=manual)	32	0.4	0.5	0	1
Number of forward gears	32	3.7	0.7	3	5
Number of carburetors	32	2.8	1.6	1	8

Descriptive statistics: in text format, selected variables

```
mydata <- mtcars
install.packages("stargazer") #Use this to install it, do this only once
library(stargazer)
stargazer(mydata[c("mpg","hp","drat")], type = "text",
          title="Descriptive statistics/selected variables", digits=1, out="table2.txt")
```

Descriptive statistics/selected variables

```
=====
Statistic N Mean St. Dev. Min Max
mpg      32 20.1   6.0   10.4 33.9
hp       32 146.7  68.6   52  335
drat     32  3.6   0.5   2.8  4.9
-----
```

The table will be saved in the working directory with whatever name you write in the `out` option. You can open this file with any word processor

```
#same output transposed and with labels instead of variable names
stargazer(mydata[c("mpg","hp","drat")], type = "text",
          title="Descriptive statistics/selected variables", digits=1, out="table2.txt", flip=TRUE,
          covariate.labels=c("Miles/(US)gallon","Gross horsepower","Rear axle ratio"))
```

Use this option if you want the variables in columns

Descriptive statistics/selected variables

```
=====
Statistic Miles/(US)gallon Gross horsepower Rear axle ratio
N          32                32                32
Mean       20.1              146.7             3.6
St. Dev.   6.0               68.6             0.5
Min        10.4              52                2.8
Max        33.9              335               4.9
-----
```

Use the option `covariate.labels` to replace variable names with variable labels. Must be in same order as in the dataset.

Descriptive statistics: in text format, selected variables, and by group

```
mydata <- mtcars
install.packages("stargazer") #Use this to install it, do this only once
library(stargazer)
# Descriptive statistics for cars with automatic transmission
stargazer(subset(mydata[c("mpg","hp","drat")], mydata$am==0),
           title="Automatic transmission", type = "text", digits=1, out="table3.txt")
```

Use subset () to
select the category

```
Automatic transmission
=====
Statistic N   Mean  St. Dev.  Min  Max
mpg       19  17.1    3.8     10.4 24.4
hp        19 160.3   53.9     62   245
drat      19   3.3    0.4     2.8  3.9
-----
```

The table will be saved in the working
directory with whatever name you
write in the out option. You can open
this file with any word processor

```
# Descriptive statistics for cars with manual transmission
stargazer(subset(mydata[c("mpg","hp","drat")], mydata$am==1),
           title="Manual transmission", type = "text", digits=1, out="table4.txt")
```

Use subset () to
select the category

```
Manual transmission
=====
Statistic N   Mean  St. Dev.  Min  Max
mpg       13  24.4    6.2     15.0 33.9
hp        13 126.8   84.1     52   335
drat      13   4.0    0.4     3.5  4.9
-----
```

Regression models: in text format

```

mydata$fast <- as.numeric((mydata$mpg > 20.1)) #Creating a dummy variable 1 = fast car
m1 <- lm(mpg ~ hp, data=mydata)
m2 <- lm(mpg ~ hp + drat, data=mydata)
m3 <- lm(mpg ~ hp + drat + factor(gear), data=mydata)
m4 <- glm(fast ~ hp + drat + am, family=binomial(link="logit"), data=mydata)
stargazer(m1, m2, m3, m4, type="text",
  dep.var.labels=c("Miles/(US) gallon","Fast car (=1)"),
  covariate.labels=c("Gross horsepower","Rear axle ratio","Four foward gears",
    "Five forward gears","Type of transmission (manual=1)"), out="models.txt")

```

The table will be saved in the working directory with whatever name you write in the out option. You can open this file with any word processor

For the output, you have the option to use variable labels instead of variable names (according to the type of model)

For the predictors, you have the option to use variable labels instead of variable names (in order they appear)

	Dependent variable:			
		Miles/(US) gallon		Fast car (=1)
	(1)	OLS	(3)	logistic
		(2)		(4)
Gross horsepower	-0.068*** (0.010)	-0.052*** (0.009)	-0.064*** (0.011)	-0.397 (1.358)
Rear axle ratio		4.698*** (1.192)	3.510* (1.851)	4.248 (21.106)
Four foward gears			-0.276 (2.135)	
Five forward gears			3.761* (2.161)	
Type of transmission (manual=1)				11.743 (359.486)
Constant	30.099*** (1.634)	10.790** (5.078)	16.306** (6.429)	29.882 (85.238)
Observations	32	32	32	32
R2	0.602	0.741	0.782	
Adjusted R2	0.589	0.723	0.749	
Log Likelihood				-1.953
Akaike Inf. Crit.				11.906
Residual Std. Error	3.863 (df = 30)	3.170 (df = 29)	3.017 (df = 27)	
F Statistic	45.460*** (df = 1; 30)	41.522*** (df = 2; 29)	24.179*** (df = 4; 27)	

Note: *p<0.1; **p<0.05; ***p<0.01

For more details/options type
?stargazer

For a nice presentation use the “html” option, open it with any word processor

OUTPUT IN HTML FORMAT

Regression models: in html format

```
mydata$fast <- as.numeric((mydata$mpg > 20.1)) #Creating a dummy variable 1 = fast car
m1 <- lm(mpg ~ hp, data=mydata)
m2 <- lm(mpg ~ hp + drat, data=mydata)
m3 <- lm(mpg ~ hp + drat + factor(gear), data=mydata)
m4 <- glm(fast ~ hp + drat + am, family=binomial(link="logit"), data=mydata)
stargazer(m1, m2, m3, m4, type="html",
  dep.var.labels=c("Miles/(US) gallon","Fast car (=1)"),
  covariate.labels=c("Gross horsepower","Rear axle ratio","Four foward gears",
    "Five forward gears","Type of transmission (manual=1)"), out="models.htm")
```

The table will be saved in the working directory with whatever name you write in the `out` option. You can open this file with any word processor

For the output, you have the option to use variable labels instead of variable names (according to the type of model)

For the predictors, you have the option to use variable labels instead of variable names (in order they appear)

In the `type` option write `html` to export R results to **html**. It may be a good idea to use the appropriate extension in the `out` option, in this example the results will be saved in the file `models.htm`.

Word can easily read `*.htm` files, making tables easily editable. Files should look like the example shown here.

Same apply to the other procedures described in the previous section.

	<i>Dependent variable:</i>			
		Miles/(US) gallon		Fast car (=1)
	(1)	OLS (2)	(3)	logistic (4)
Gross horsepower	-0.068*** (0.010)	-0.052*** (0.009)	-0.064*** (0.011)	-0.397 (1.358)
Rear axle ratio		4.698*** (1.192)	3.510* (1.851)	4.248 (21.106)
Four foward gears			-0.276 (2.135)	
Five forward gears			3.761* (2.161)	
Type of transmission (manual=1)				11.743 (359.486)
Constant	30.099*** (1.634)	10.790** (5.078)	16.306** (6.429)	29.882 (85.238)
Observations	32	32	32	32
R ²	0.602	0.741	0.782	
Adjusted R ²	0.589	0.723	0.749	
Log Likelihood				-1.953
Akaike Inf. Crit.				11.906
Residual Std. Error	3.863 (df = 30)	3.170 (df = 29)	3.017 (df = 27)	
F Statistic	45.460*** (df = 1; 30)	41.522*** (df = 2; 29)	24.179*** (df = 4; 27)	

Note:

*p<0.05 **p<0.01 ***p<0.001